

EXHIBIT 7

**UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF TEXAS
SHERMAN DIVISION**

THE STATE OF TEXAS et al.

§

Plaintiffs,

§

Civil Action No 4:20-cv-00957

vs.

§

GOOGLE LLC,

§

Defendant.

§

JURY TRIAL DEMANDED

§

§

§

PLAINTIFFS' FIRST SET OF REQUESTS FOR PRODUCTION (NO. 1)

In accordance with Federal Rules of Civil Procedure 26 and 34, and Eastern District of Texas Local Rule CV-26, Plaintiffs hereby request that Defendant produce documents to Shawn Cowles and David Ashton, Texas Office of the Attorney General, 300 W. 15th Street, 7th Floor, Austin, Texas 78701, or if delivered by mail, at P.O. Box 12548, Austin, Texas 78711, as specified below. Documents produced by Defendant must adhere to the Definitions and Instructions set forth below, the conditions of Rule 34 of the Federal Rules of Civil Procedure, and Eastern District of Texas Local Rule CV-26. The documents and things to be produced are described as follows:

Definitions

As used herein:

1. **“Document”** shall mean the original and all non-identical copies (whether different from the original because of notes, underlining, attachments, or otherwise), all computer files, and any written, printed, or recorded material of every kind in the possession, custody, or control of You (defined below). The term “computer files” includes information stored in or accessible through computers, portable computers, handheld devices, workstations, minicomputers, mainframes, servers, backup disks and tapes, archive disks and tapes, and other forms of offline storage, whether on or off the premises of Your place of business. This definition covers electronic mail messages (“e-mail”), wikis, slack messages, text messages, social media messages, and other electronic Documents in Your possession, custody, or control.

2. **“Identify”** with respect to a Document means to give, to the extent known, the type of Document, the general subject matter, the date of the Document, and the author, addressees, and recipients of the Document.

3. **“Identify”** with respect to a Person means to give, to the extent known, the Person’s full name, title, and when referring to a natural Person, additionally, the present or last known place of employment. Once a Person has been first identified in accordance with this definition, only the name of that Person need be listed in response to subsequent discovery requesting the identity of that Person.

4. **“Native”** or **“Native Format”** for an electronic document shall mean the format in which it is ordinary stored as a separate file. In the case of documents that are commonly stored online or “in the cloud,” Native Format means the format with the nearest functionality to its online equivalent—not necessarily the format with the nearest appearance. *See, e.g.,*

https://support.google.com/vault/answer/6099459#drive_export. Unless a file is ordinarily stored as a PDF, extracting or exporting this file as a PDF is not its native format.

5. **“OAG”** means the Office of the Attorney General of Texas.
6. **“Person”** is defined as set forth in TEX. BUS. & COM. CODE § 15.03(3), and includes any natural person, corporate entity, partnership, association, joint venture, government entity, or trust.
7. **“You”, “Your”, “Alphabet”, or “Google”** shall mean Alphabet Inc., its past and present officers, employees, agents and representatives, parents and predecessors, divisions, subsidiaries, affiliates, partnerships and joint ventures. The terms “subsidiary”, “affiliate”, and “joint venture” refer to any firm in which there is total or partial ownership of twenty-five percent or more or total or partial control between the company and any other Person or entity.

Instructions

1. To the extent that documents responsive to one of these requests has already been produced, identify any such documents by BATES label.
2. These requests are continuing in nature and should be updated regularly.
3. Documents should be produced in the format described in Attachment A to these requests. Please be advised that Attachment A may not be the same as any similar attachment to a prior discovery request or civil investigative demand.
4. Pursuant to Rule 34(b)(2)(D), produce all Documents in their Native Form. To the extent that a Document has otherwise been produced to Plaintiff in a different form, produce all native versions with a label corresponding to the first page, if any, of documents produced in another form. For example, if a document was produced with four images using BATES labels EXAMPLE_0000007-0000010, the native version would have the BATES label EXAMPLE_0000007. The native, if any, for the subsequent document would be labeled EXAMPLE_0000011.
5. Explain in detail the process used to identify and collect documents responsive to each request for production. Your explanation should include *inter alia* the extent to which search terms were utilized, if at all, and a general description of the collection of documents to which the search terms were applied.
6. If You contend that a request in this document encompasses another request in this document or a prior document, please identify the request or requests that are allegedly encompassed. However, produce all documents responsive to said request that are not otherwise encompassed by said other request. For example, if You contend that all documents responsive to RFP 25 are also responsive to RFP 30, state that this is Your contention, produce and identify

documents that are responsive to RFP 25 and also produce and identify documents that are responsive to RFP 30 that are not also responsive to RFP 25. This instruction does not modify Google's obligation to produce all responsive documents to each request for production.

7. If You contend that a request in this document is encompassed by a request in this document or a prior document, please identify any such other requests and further identify by BATES label the documents that You contend are only responsive to the allegedly encompassed request. For example, if You contend that RFP 26 is encompassed by RFP 14, identify the documents that You contend are responsive to RFP 26. This instruction does not modify Google's obligation to produce all responsive documents to each request for production.

8. If all or any portion of any responsive document is withheld for any reason, including a claim of attorney-client privilege, any other applicable privilege or immunity, or judicial order, please submit, within the time permitted for a response, a list of the documents withheld and state individually as to each document: (1) the source of the document; (2) the name, address, position and organization of the author and recipient of the document; (3) the type, title, specific matter, length and date of the document; and (4) the legal or factual basis for withholding the document, including, but not limited to, a general description of the redacted matter, and an identification of the actual or anticipated litigation for a claim based on work product immunity.

Request

REQUEST FOR PRODUCTION NO. 1

Produce all documents that were produced to the OAG bearing BATES stamps GOOG-TEX-00000001 to GOOG-TEX-01279945 in native form.

Dated: April 1, 2021

/s/ Ashley Keller
Ashley Keller
Admitted Pro Hac Vice
ack@kellerlenkner.com
Brooke Smith
brooke.smith@kellerlenkner.com
150 N. Riverside Plaza, Suite 4270
Chicago, Illinois 60606
(312) 741-5220
Warren Postman
wdp@kellerlenkner.com
1300 I Street, N.W., Suite 400E
Washington, D.C. 20005
(202) 749-8334
KELLER LENKNER LLC

/s/ Mark Lanier
Mark Lanier (*lead counsel*)
Texas Bar No. 11934600
Mark.Lanier@LanierLawFirm.com
Alex J. Brown
Alex.Brown@LanierLawFirm.com
Zeke DeRose III
Zeke.DeRose@LanierLawFirm.com
10940 W. Sam Houston Parkway N. Suite 100
Houston, Texas 77064
Telephone: (713) 659-5200
Facsimile: (713) 659-2204
THE LANIER LAW FIRM, P.C.

Attorneys for Plaintiff States of Texas, Idaho, South Dakota, and North Dakota

On behalf of all Plaintiff States

ATTACHMENT A

DATA DELIVERY PRODUCTION REQUIREMENTS

Prior to any production, a meeting with the Antitrust Division is required to discuss the location of any potentially relevant Electronically Stored Information (ESI) – to include discussing E-mail, Network Shares, Local Storage, Databases, Mobile Devices, the Cloud, Chat Applications, and Social Networking.

This document describes the standard specifications and procedures for making an image-based production to the Antitrust Division of the Office of the Attorney General of Texas (“OAG”) in the form of Load Files.

- To ensure the efficient processing and review of any electronic production, legal, technical, and economic OAG staff need to resolve the details prior to production, and preferably before you or your vendor begin to gather and process responsive documents.
- Care should be taken to ensure that all responsive data and metadata are preserved in the collection process.

I. Categories of Documents

Discussion regarding the details of an electronic production should focus on seven categories of documents:

- A. Electronic documents
- B. Email and other electronic messages (e.g. instant messaging and chat applications)
- C. Databases
- D. Hard copy documents
- E. Audio and video data
- F. Foreign language materials

General requirements for each category of document are outlined below. For information regarding document-specific metadata and bibliographic information, please refer to the enclosed Metadata Table of Requested Fields.

II. General Instructions for Productions

To ensure the efficient processing and review of any electronic production, contact the OAG before you or your vendor begin to gather and process responsive documents.

- A. A cover letter should be included with each production of data. The following information should be included in the letter:
 1. List of each piece of media (hard drive, thumb drive, DVD or CD) included in the production by the unique number assigned to it, and readily apparent on the physical media.
 2. List of fields in the order in which they are listed in the data file
 3. Time zone in which emails were standardized during conversion (email collections only). Label all media with the following:

- a. Case number
 - b. Volume Name
 - c. Production date
 - d. Bates range
 - e. Media number (1 of X), if applicable
- B. Documents created or stored electronically MUST be produced in their original electronic format to the maximum extent possible using the minimum number of format conversions. For example, word-processing documents should not be printed to paper or PDF with those versions being used to create supplemental image files. In this example, the word-processing documents should be produced as-is, and any images created directly from the original format.
- C. To the extent possible, all times should be standardized to UTC.
- D. To the extent that an electronic document has an owner, this owner should be identified as the custodian in that document's metadata regardless of the number of people that may have access to a document. The owner does not need to be a natural person and may refer to a group, organizational unit, or an entire company, as appropriate.
- E. Data should be produced using whichever media requires the least number of deliverables.
- F. All productions should be checked and produced free of computer viruses.
- G. Password-Protected Files: to the extent any produced documents are password-protected, the Producing Party must either unlock the document prior to production or provide passwords in order to allow access by the Receiving Party. If the Producing Party is unable to process a document because of unknown passwords or encryption, the Parties shall meet and confer to determine the appropriate steps to take with respect to such document.
- H. Production Load Files may also be transmitted electronically (*e.g.*, file share sites, FTP). Load Files in excess of 5GB should be discussed with the OAG in advance of transmission.

III. Categories of Documents

A. Electronic Documents

Electronic documents include word-processing documents, spreadsheets, presentations, file system directories, and all other electronic documents not specifically discussed elsewhere. To the extent possible or otherwise indicated, electronic documents should be produced as native files

in the format in which they are kept in the usual course of business. For specific types of Google documents, the following are considered to be native formats¹:

- | | |
|--------------------|-------|
| a. Google Docs | .DOCX |
| b. Google Sheets | .XLSX |
| c. Google Slides | .PPTX |
| d. Google Drawings | .PDF |

All passwords and encryption must be removed from electronic documents prior to production. All metadata and bibliographic information for each electronic document must be extracted based on the manner in which the documents are usually kept. Refer to the Metadata Table of Requested fields for required metadata fields that must be extracted. The native file should be named per the BEGBATES# with the appropriate file extension.

For electronic documents that are capable of being rendered in image format, single-page TIFF images should also be provided. The first TIFF image should be stamped with the value of the BEGBATES# and subsequent images should be stamped sequentially. The special image instructions are described below.

1. Spreadsheets

For each spreadsheet, provide a single placeholder image for the spreadsheet. The placeholder image must contain at a minimum the BEGDOC#, FILENAME, and FILEPATH. The Bates range for a spreadsheet should be a single number (*e.g.*, ABC-JD-00000001 – ABC-JD-00000001).

2. Presentations

For each presentation, its images should be produced in full slide image format along with speaker notes (which should follow the full images of the slides). The linked native file name should match the BEGDOC#, FILENAME, and FILEPATH.

3. Hidden Text

All hidden text (*e.g.*, track changes, hidden columns, hidden slides, mark-ups, notes) shall be expanded and rendered in the associated image files.

4. Embedded Files

All embedded objects (*e.g.*, graphical files, Word documents, Excel spreadsheets, .wav files) that are found within an electronic document shall be produced so as to maintain the integrity of the source document as a single document. For purposes of production, the embedded files shall remain embedded as part of the original source document. Internal documents that are embedded

¹ The listed extensions are merely common examples of file extensions associated with those formats and are not to be construed as limitations. In other words, Word documents with extensions other than .DOC and .DOCX should also be produced in native format.

or referenced via hyperlinks and hyperlinked shared resources must be produced as separate, attached documents following the instructions in this appendix.

5. Image-Only Files

All image-only files (non-searchable PDFs, multi-page TIFFs, Snipping Tool screenshots, etc., as well as all other images that contain text) shall be produced with associated OCR text, metadata, and identifying information.

6. Archive File Types

Extract all applicable metadata for archive files (*e.g.*, .zip, .rar, .tgz) as for all other files. For the archive file itself, provide a single placeholder image. The placeholder image must contain at a minimum the BEGDOC#, FILENAME, and FILEPATH. The files within the archive must be uncompressed and produced using the appropriate methodology for the file's type. Each file contained within an archive file should be produced as a child to the parent archive file. If the archive file is itself an attachment, that parent/child relationship must also be preserved.

7. Directories

Extract all applicable metadata for directory files as for all other files. For the directory itself, provide a single placeholder image. The placeholder image must contain at a minimum the BEGDOC#, FILENAME, and FILEPATH. To the extent that the files associated with the directory are being produced, they should be produced as child documents of this directory.

8. Other File Types and Non-PC or Non-Windows Based Systems

For all other file types, extract all applicable metadata, and provide a single placeholder image for the electronic document. The placeholder image must contain at a minimum the BEGDOC#, FILENAME, and FILEPATH. If you contend that this is inappropriate for a particular file type, please contact the OAG to determine the optimal format of production.

B. Email, Attachments, and Other Electronic Messages

Email and other electronic messages (*e.g.*, instant messages (IMs)) should be produced in an appropriate native format. In addition and to the extent possible, electronic messages should be rendered as image files with related searchable text and metadata. If the company identifies message participants in a form other than name (*e.g.*, email address, employee #), a table matching these identifiers to names must be provided.

Google Chat and Google Hangout archives should be produced as one or more MBOX files, if possible. Otherwise, they should be produced using the appropriate number of JSON files contained within individual zip files for each conversation. Slack channel conversations should be exported as the appropriate number of JSON files contained within individual zip files for each conversation.

Email repositories, also known as email databases (*e.g.*, Outlook .PST, Lotus .NSF), can contain a variety of items, including messages, calendars, contacts, tasks, etc. For purposes of production, responsive items should include the "Email" metadata/database fields outlined in the Metadata Table, including but not limited to, all parent items (mail, calendar, contacts, tasks, notes,

etc.) and child files (attachments of files to email or other items), with the parent/child relationship preserved. Similar items found and collected outside an email repository (*e.g.*, .MSG, .EML, .HTM, .MHT) should be produced in the same manner.

Each IM conversation should be produced as one document.

Outlook (.PST) and Lotus Notes (.NSF) email files may be produced in native file format. Production of these native files should be discussed in advance with the OAG.

Electronic phone records should be discussed with the OAG in advance of production to determine the optimal format of production.

C. Database Productions

Production of enterprise databases are not addressed in these specifications and must be discussed with the appropriate legal and technical staff to determine the optimal production format; these will usually fall outside the scope of an image-based production. Care must be taken to ensure that all responsive databases and their metadata are preserved.

Data extracted from a database must be produced in CSV format. The first row of the file must identify each field of the records contained in the file. Provide a single placeholder image. The placeholder image must contain at a minimum the BEGDOC#, FILENAME, FILEPATH, and the query used to generate the data contained in the CSV file. (*e.g.*, “SELECT field1, field2 FROM Table...”).

D. Hard-Copy Documents

Hard-copy documents (*e.g.*, documents ordinarily maintained in paper) are to be produced as black-and-white image files, except where noted below, with related searchable OCR text and identifying information. Special attention should be paid to ensure that hard-copy documents are produced as they are kept, reflecting attachment relationships between documents and information about the file folders within which each document is found. An image of the file folder or an appropriate placeholder image should be produced for any such file folders. In addition, multi-page documents must be produced as single documents (*i.e.*, properly unitized) and not as several single-page documents. Where color is required to interpret the document, such as hard copy photos, and certain charts, its image MUST be produced in color. These color images are to be produced as JPEG format. Hard-copy photographs should be produced as color JPEGs, if originally in color, or grayscale TIFF files if originally in black-and-white.

E. Audio and Video Data

Audio and video file types not specifically discussed elsewhere should be discussed with the OAG in advance of production to determine the optimal format of production.

F. Foreign Language Materials

Production of Foreign Language Materials not addressed in these specifications must be discussed with the appropriate legal and technical staff to determine the optimal production format prior to production.

IV. Processing

A. De-Duplication

Before doing any de-duplication, provide the OAG with a written description of the method that will be used to de-duplicate (including which elements are compared and what hash codes are used), and what is considered a duplicate. Then confirm that your approach is acceptable to the OAG. De-duplication of hard-copy documents, or that of “loose” electronic documents (*e.g.*, presentation slides located on the custodian’s C: drive) against email attachment versions of those same documents, is unacceptable. The integrity of any produced email and any related “document family” must be maintained except as limited by any claim of privilege. Vertical de-duplication is crucial when a production includes electronic documents from back-ups. Horizontal de-duplication must be done in a way that preserves (and produces) information on blind copy (Bcc) recipients of emails and other custodians whose files contain the duplicates that will be eliminated from the production.

A Custodian Append File is to be produced when de-duplicating ACROSS custodians (*i.e.*, horizontal de-duplication) and data is produced on a rolling basis. The file must be provided on an incremental basis **starting with the second submission**; as more custodians are discovered for previously produced documents, this file is updated with **only the new** custodian information.

The Custodian Append File is a two-field delimited file consisting of the BEGDOC#s of the previously delivered document and the **new** custodian names for the duplicates of those records that would otherwise have been produced in the subsequent (new) submissions.

These specifications do not allow for near de-duplication or email threading. These formats must be discussed separately with the OAG and consent obtained prior to the use of such techniques for production.

B. Document Numbering

Documents must be uniquely and sequentially Bates-numbered across the entire production, with an endorsement burned into each image. Each Bates number shall be of a consistent length, include leading zeros in the number, and unique for each produced page. Bates numbers should contain **no more** than three segments. For example, a company identifier, a middle segment identifying the custodian, and a sequential page counter with connecting hyphens. The number of digits in the numeric portion of the format should not change in subsequent productions, nor should spaces, hyphens, or other separators be added or deleted. Under no circumstances should bates numbers contain embedded spaces, slashes (/), backslashes (\), carats (^), ampersands (&), hash marks (#), plus signs (+), percent signs (%), dollar signs (\$), exclamation marks (!), pipes (|), any character used as a delimiter in the metadata load files, or any character not allowed in Windows file-naming convention (,\ / : * ? “<> | ~ @ ^). Bates numbers may contain hyphens (-).

C. Privilege Designations

Documents redacted pursuant to any claim of privilege will be designated “Redacted” in the SPECPROPERTIES field as described in the Metadata Table. Appropriately redacted searchable text (OCR of the redacted images is acceptable), metadata, and bibliographic

information must also be provided. All documents that are part of a document family that includes a document withheld pursuant to any claim of privilege will be designated “Family Member of Privileged Doc” in the SPECPROPERTIES field as described in the Metadata Fields table for all other documents in its family. Placeholder images with BEGDOC#, FILENAME, FILEPATH and reason withheld (*e.g.*, “Privileged”) should be provided in place of the document images of the privileged document.

D. Redactions

To the extent that a document will be produced with redactions based on an assertion of privilege, the redaction should contain a brief explanation of the basis for the redaction, if not readily apparent based on context. For example, if a redaction were immediately preceded by “SSN”, it would be clear that a social security number was being redacted. At a minimum, the redaction must indicate whether the alleged privilege is attorney/client, work-product, or other privilege. For example, a redaction might include the text “WP” to indicate that the underlying material is being withheld on the basis of being work-product. To the extent possible, the rationale shall be indicated on the redaction itself. Otherwise, a chart should accompany the production of the redacted document. To the extent that an email is produced in redacted format, all recipients of the allegedly privileged email must be identified, including for any forwards of that email or replies to that email.

E. Metadata

The metadata for an electronic document includes both the metadata within the document itself as well as the corresponding entry in a file system, if any. In other words, a document’s filename is distinct from a title that may be embedded in the document.

IMAGE AND TEXT FILE SPECIFICATIONS, AND LOAD FILE CONFIGURATION

I. Load File Configuration

- Each production must have a unique VOLUMENAME (e.g., TXAG_VOL001)
- The VOLUMENAME should increase sequentially with each subsequent production (e.g., TXAG_VOL001, TXAG_VOL002, etc.)
- Under the VOLUMENAME folder, the production should be organized into 4 subfolders:
 - DATA (should contain metadata, DAT, OPT, custodian append, etc. files)
 - IMAGES (may contain subfolders, with no more than 5,000 images per folder)
 - NATIVES (may contain subfolders, containing linked native files, with no more than 5,000 files per folder)
 - TEXT (may contain subfolders, with document-level text files)

II. Image/Native File Specifications

- Black-and-white Group IV Single-Page TIFFs (300 DPI). Color images should be provided in jpg format
- Image file names should match the page identifier for that specific image and end with the .tif (or .jpg if needed) extension.
 - Example: ACME-ABC-0003072.TIF
- File names cannot have embedded spaces, commas, ampersands, slashes, back slashes, hash marks, plus signs, percent signs, exclamation marks, any character used as a delimiter in the metadata load files, or any character not allowed in Windows file-naming convention. (, & \ / # + % ! : * ? “ < > | ~ @ ^)
- Images for a given document must reside together in the same folder.
- The maximum number of image files should be limited to 5,000 per folder.
- Native file names should match the BEGDOC entry for that specific record and end with the appropriate file extension.
- The maximum number of native files in a subfolder should be limited to 5,000 per folder.
- Any encryption or password protection must be removed from all native format files produced.

III. Extracted or Searchable Text File Specifications and Control List Configuration

- Extracted text, as opposed to OCR, must be provided with all records, except for documents that originated as hard copy or redacted documents.
 - For hard copy documents, provide OCR text.
 - For redacted documents, provide OCR text for the redacted version.
- There should be a single extracted/OCR text file per document, in ASCII Text format only.
- The name of the text file should be the same as the document's first page/Bates number, with a TXT extension: BEGDOC#.TXT.
- There must be a carriage return and line feed (CRLF) no later than the 250th character of the first line of every text file.

- All soft and hard returns in the native electronic or image file should be replicated as a Carriage Return Line Feed (CRLF) in the text file (i.e. the lines of text in the file terminate with a CRLF in correlation with the appearance of the native electronic or rendered image file). Pay particular attention to not allow multi-line paragraphs of e-mails to be rendered as a single, extended line of text.
- Text files should include page breaks that correspond to the “pagination” of the image files. Place text files under a “TEXT” folder and provide a Control List file for loading in the “DATA” folder on the delivery media.

IV. Image Load File (.opt) Configuration

- Page level comma-delimited file containing seven fields per line.
 - PageID,VolumeLabel,ImagePath,DocumentBreak,FolderBreak,BoxBreak,PageCount
 - PageID – PageID of the item being loaded. MUST be identical to the image name (less the file extension).
 - VolumeLabel – Optional. If used it is preferable that it match the VOLUMENAME assigned in the corresponding metadata load file.
 - ImageFilePath – The path to the image from the root of the delivery media.
 - DocumentBreak – The letter “Y” denotes the first page of a document. If this field is blank the page is not the first page of a document.
 - FolderBreak – Leave empty
 - BoxBreak – Leave empty
 - PageCount – Optional
- Example - ABC-JD00030005,,\ABC002\Images\001\ ABC-JD-00030005.tif,Y,,,

V. Metadata Load File Delimiters and Configuration

- Field Separator ¶ (ASCII 020)
- Text Qualifier ¢ (ASCII 254)
- Substitute Carriage Return or New Line in data ® (ASCII 174)
- Multi-value separator (Do Not Follow with Space) ; (ASCII 059)
- Date format YYYYMMDD (date type fields only)
- Time format HH:MM:SS in 24-hour format (e.g., 04:32 pm formatted to 16:32:00 – Do not include AM, PM, or timezone² indicators).
- There should be one line for every record in the load file. A carriage return and line feed (CRLF) must appear at the end of each record and ONLY at the end of each record.
- The first row of each metadata load file should be a header row containing the field names. Field names must match OAG Metadata Table field names.
- All requested fields should be present in the metadata load file whether data exists or not. Field order must remain consistent in subsequent productions.

² Note that time zone information is still required in a different metadata field.

VI. METADATA TABLE OF REQUESTED FIELDS

Field Name	Field Description	Field Type
VOLUMENAME	Production volume number (e.g., ABC001-001)	Note Text
BEGDOC	Start Bates (including prefix) -- No spaces or special characters	Note Text
ENDDOC	End Bates (including prefix) -- No spaces or special characters	Note Text
ATTACHRANGE	Range of the BEGDOC value of the parent record to the ENDDOC value (including prefix) of the last child record (for example, ABC-JD-00001201 ABC-JD-00001220); populated for all documents in the group. Empty if the record is NOT in family grouping	Note Text
BEGATTACH	Starting Bates of attachment range	Note Text
ENDATTACH	Ending Bates of attachment range	Note Text
PARENTBATES	Parent record's BEGDOC#, including prefix (populated ONLY in child records)	Note Text
CHILDBATES	Child document list: BEGDOC# of each child, separated by semicolons [;] (populated ONLY in parent records)	Multi-Entry
NUMPAGES	Page count	Integer
DOCTYPE	Examples: Email, Presentation, Spreadsheet, Directory	
SPECPROPERTIES	Indicate all that apply : Record Type: E-Doc, E-Doc Attachment, Email, Email Attachment, Hard Copy, Calendar Appt Other Notations: Translation of [DOCID of original], Translated as [DOCID of Translation] Privilege Notations: Redacted, Privileged, Family Member of Priv Doc	Multi-Entry
CUSTODIAN	Custodian(s) / source(s) - format: Last, First or ABC Dept For files stored using Google Drive, the CUSTODIAN field corresponds to the "owner" field. (see below)	Multi-Entry
COLLABORATORS	(see below)	Multi-Entry
VIEWERS	(see below)	Multi-Entry

Field Name	Field Description	Field Type
OTHERS	(see below)	Multi-Entry
FILEID	A unique identifier for the electronic document as reported by the filesystem.	Note Text
PARENTID	A unique identifier for the parent of the electronic document or directory.	Note Text
DRIVEID	A unique identifier for the shared drive. For files stored using Google Drive, the DRIVEID field corresponds to the “driveId” field.	Note Text
TITLE	The name or title of the document as reflected in the document’s metadata.	Note Text
TIMEZONE	The TimeZone from which the native file was collected.	Note Text
FROM	Author of the Email or Calendar item (as formatted on the original)	Note Text
TO	Recipients of the Email (as formatted on the original)	Multi-Entry
CC	Names of the individuals who were copied on the Email (as formatted on the original)	Multi-Entry
BCC	Names of the individuals who were blind-copied on the Email (as formatted on the original)	Multi-Entry
SUBJECT	Email, document, or calendar subject	Note Text
DATESENT	Date the Email was sent. Format: YYYYMMDD.	Date
TIMESENT	Time Email was sent -- Format: HH:MM:SS (use 24 hour times, e.g., 13:32 for 1:32 pm; time zone indicators cannot be included)	Time
DATERECEIVED	Date Email was received. Format: YYYYMMDD.	Date
TIMERECEIVED	Time Email was received. Format: HH:MM:SS (use 24 hour times, e.g., 13:32 for 1:32 pm; time zone indicators cannot be included)	Time
HEADER	The internet header information for Email sent through the internet;	Note Text
INTERNETMSGID	Globally unique identifier for a message which typically includes messageid and a domain name. Example: <0E6648D558F338179524D555@m1p.innovy.net	Note Text
MESSAGEID	Proprietary email database/mailstore/post office file associated with centrally managed enterprise email servers. Microsoft Outlook PST EntryID, the UniqueID (UNID) for Lotus Notes, equivalent value for other proprietary mailstore formats.	Note Text
CONVERSATIONINDEX	Email Thread Identification	Note Text

Field Name	Field Description	Field Type
IMPORTANCE	Email flag indicating priority level set for message	Note Text
DATECREATED	Date file was created. Format: YYYYMMDD.	Date
REVISION	Revision number extracted from metadata of electronic documents	Note Text
DATESAVED	Date file was last modified. Format: YYYYMMDD.	Date
ORGANIZATION	Company field extracted from the metadata of an electronic document	Note Text
KEYWORDS	List of keywords, tags, and categories extracted from the metadata of an electronic document	
AUTHOR	Author field value extracted from the metadata of the electronic document	Note Text
SPEC#	Request for production number(s) to which the document is responsive	Multi-Entry
HASHMD5	Document MD5 hash value (used for de-duplication or other processing) NOTE: only one hash value metadata field need be provided—either MD5 hash or SHA1 hash.	Note Text
HASHSHA	Document SHA1 hash value (used for de-duplication or other processing) NOTE: only one hash value metadata field need be provided—either MD5 hash or SHA1 hash.	Note Text
FILESIZE	File size in Bytes (integer value only - do not include unit of measure or decimal places - e.g., 568)	Integer
FILENAME	File name of electronic document with the file extension (e.g., Re: tomorrow sounds fun.eml; Read this today.docx)	Note Text
FILEEXTENSION	File extension of electronic document (e.g., .eml, .msg, .xlsx)	Fixed Length 5 chars
FOLDERLABEL	Email folder path (sample: Inbox\Active); or Hard Copy folder/binder title/label	Note Text
FILEPATH	File path to electronic document as it existed in its original environment (see below)	Note Text
NATIVELINK	File path location to the current native file location on the delivery medium	Note Text

NOTES REGARDING METADATA FIELDS

CUSTODIAN

- For emails, the CUSTODIAN field should contain the mailbox that contained the email without regard to whether the mailbox's owner is an identified party to the email, such as a BCC or distribution list recipient

- For electronic directories, the CUSTODIAN field is the owner of the directory or drive identified by the containing filesystem. The owner may be a group or organizational unit.
- For electronically stored files, the CUSTODIAN field is the owner of the file identified by the containing filesystem. If ownership of a file is determined by ownership of the containing directory, the CUSTODIAN field is the owner of that directory. The owner may be a group or organizational unit.
- For scanned documents, the CUSTODIAN field should be populated with the name of the person who was in possession of the document at the time it was collected.

FILEPATH

For electronically stored files, the FILEPATH field should be expressed in a human-readable form in the manner that the file is accessed in the ordinary course of business. This field should be populated with the names of all folders or directories between the root folder and the particular electronically stored document. For example, “C:\Folder\SubFolder\”. The FILEPATH should not reflect the path created by virtue of collecting the document for purposes of this litigation. For example, if a DOCUMENT1 is typically accessed as \\DIVISION\PROJECT\DOCUMENT1, the FILEPATH should be \\DIVISION\PROJECT\DOCUMENT1 and not \\USER1\DOCUMENT1 (assuming that it was collected with respect to a custodian designated USER1).

COLLABORATORS

This field should be populated with an identifier for each account or organizational unit that has been explicitly³ granted write-access to the applicable file or directory. The identifier for each account or organizational unit should be separated by semicolons.

VIEWERS

This field should be populated with an identifier for each account or organizational unit that has been explicitly granted read-access to the applicable file or directory but does not have write-access to the file or directory. Accounts that also have write-access should be identified only in the COLLABORATORS field. The identifier for each account or organizational unit should be separated by semicolons.

OTHERS

This field should be populated with an identifier for each account or organizational unit that has been explicitly granted a form of access to the applicable file or directory that cannot be expressed as read access or write access. An account listed as in the VIEWERS or COLLABORATORS field can also be listed in the OTHERS field. The identifier for each account or organizational unit should be separated by semicolons.

³ In this context, “explicitly” means identified within a corresponding Access Control List (ACL). For example, if an organizational unit is identified in the ACL, the corresponding field would contain an identifier for the organizational unit, but not for the accounts corresponding to members of that organizational unit unless those members (or accounts) are independently identified in the ACL.

